

Student Name: Tony Yang
Student Number: 1003289229

Introduction

This report aims to analyze patient data stratified by dementia status, encompassing individuals diagnosed with dementia, those without, and those transitioning to dementia over time. The primary focus is to investigate the impact and interaction of dementia on eTIV and nWBV across these distinct groups and multiple visits.

1. Research Questions 1: Is there a notable variance in the estimated Total Intracranial Volume (eTIV) across different groups and visit periods and does the interaction between group affiliation and visit time yield a significant effect?

2. Research Questions 2: Is there a notable variance in the Normalize Whole Brain Volume (nWBV) across different groups and visit periods and does the interaction between group affiliation and visit time yield a significant effect?

Data Preparation

Upon initial examination, the raw dataset contains 16 columns and 294 rows. Following my review, I determined that no data cleaning is necessary for the scope of my analysis. Since my analysis is quantitative, I have reduced our working dataset to the following columns from the raw dataset. Below is a short description of each column:

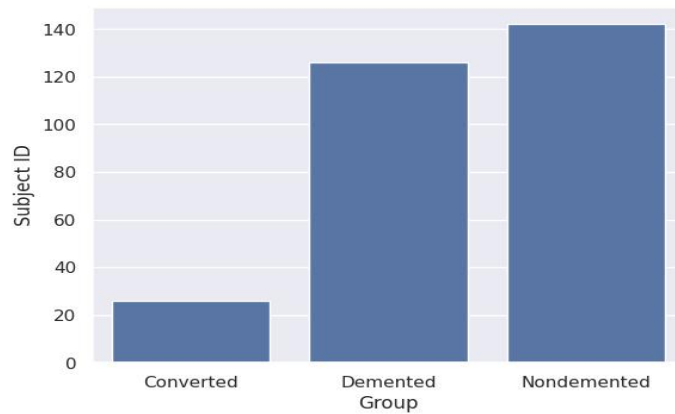
- Subject ID: Subject Identification
- Group: Group of a patient
- Visit: Visit number of a patient
- eTIV: Estimated Total Intracranial Volume
- nWBV: Normalize Whole Brain Volume

Exploratory Data Analysis (EDA)

After data preparation, I delved into a comprehensive Exploratory Data Analysis (EDA) to uncover insights from the dataset. I began by describing our numerical data, as shown in Table 1, which presents a summary of these variables.

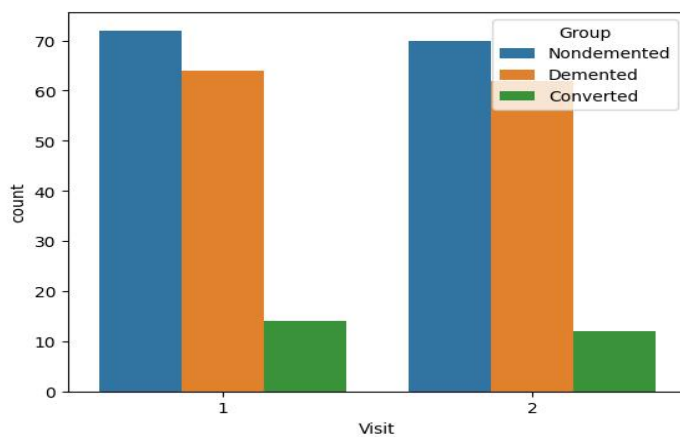
	Visit	eTIV	nWBV
count	294.00	294.00	294.00
mean	1.49	1478.85	0.73
std	0.50	176.56	0.04
min	1.00	1106.00	0.65
25%	1.00	1347.25	0.70
50%	1.00	1461.50	0.73
75%	2.00	1569.00	0.76
max	2.00	2004.00	0.84

Table 1 : Summary Statistics



I have generated a bar plot shown in Figure 1 to visually illustrate the distribution of subjects across different groups. Upon examination of the plot, it is evident that the nondemented group exhibits the highest count of subjects, followed by the demented group in second place, and the converted group with the lowest count.

Figure 1: Bar Plot For Subject ID In Different Groups



Moreover, I analyzed a bar plot illustrating the count of visit times across various groups, offering a more comprehensive perspective as depicted in Figure 2. It is evident from the plot that the Nondemented Group exhibits the highest count for both visit 1 and visit 2, followed by the Demented Group in second place, with the Converted Group having the lowest count.

Figure 2: Count Plot For Visit Among Different Groups

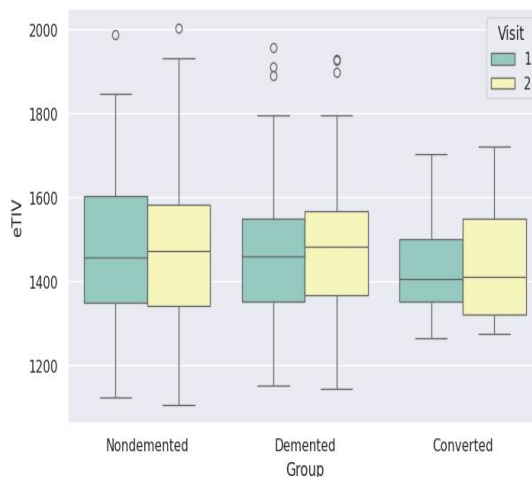


Figure 3: Box Plot for eTIV

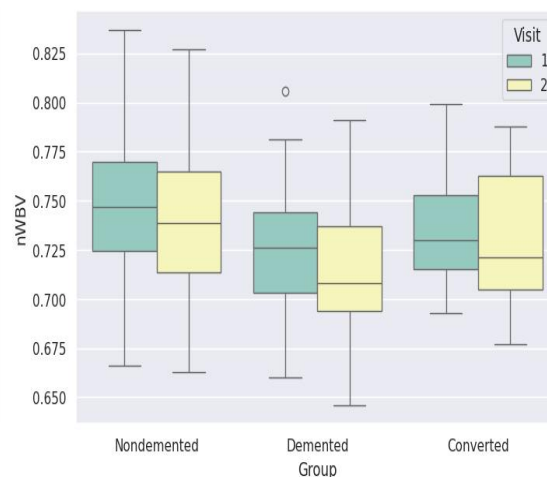


Figure 4: Box Plot for nWBV

Upon reviewing the box plots for eTIV and nWBV across different groups and visit times in Figure 3 and 4. No clear patterns emerge. While the data distribution appears somewhat normal, with similar means across the plots, further analysis of mixed-effect ANOVA is necessary to draw conclusive insights.

Mixed Effects ANOVA

Research Questions 1: Is there a notable variance in the estimated Total Intracranial Volume (eTIV) across different groups and visit periods and does the interaction between group affiliation and visit time yield a significant effect?

Hypothesis:

H0: There is no significant difference and there is no interaction effect

H1: There is a significant difference and there exists an interaction effect

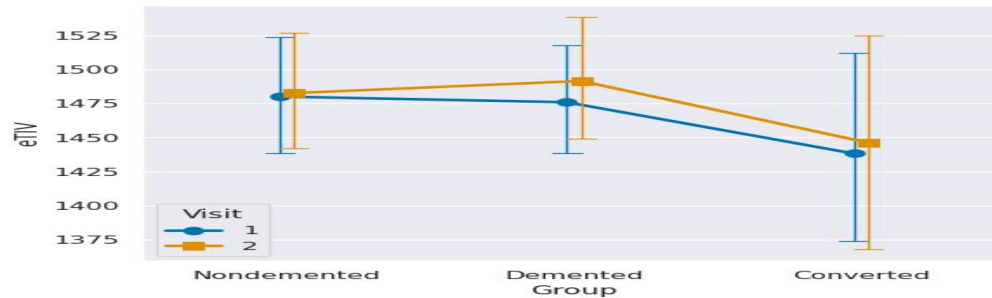


Figure 5: Point Plot of eTIV Across Different Groups and Visits

Table 2: Mixed Effect Anova Table

	Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
0	Group	37424.708	2	141	18712.354	0.297	0.743	0.004	NaN
1	Visit	5573.920	1	141	5573.920	9.225	0.003	0.061	1.0
2	Interaction	1004.783	2	141	502.392	0.831	0.438	0.012	NaN

Based on Table 2's statistical analysis, eTIV differences across groups aren't significant ($F = 0.297$, $p = 0.743 > 0.05$). However, eTIV varies significantly over visit ($F = 9.225$, $p = 0.003 < 0.05$). The interaction between group and visit doesn't significantly affect eTIV ($F = 0.831$, $p = 0.438 > 0.05$). In summary, while visit periods notably influence eTIV, neither group nor their interaction significantly affect it.

To validate the Mixed Effect ANOVA, certain assumptions need to be met.

Assumption 1: Mauchly's test: P value = 1 > 0.05 , assumption of sphericity is met.

Assumption 2: Shapiro-Wilk test: All P values > 0.05 , assumption of normality is met.

Assumption 3: Levene's test: P values > 0.05 for both within groups. assumption of homogeneity is met.

Research Questions 2: Is there a notable variance in the Normalize Whole Brain Volume (nWBV) across different groups and visit periods and does the interaction between group affiliation and visit time yield a significant effect??

Hypothesis:

H0: There is no significant difference and there is no interaction effect

H1: There is a significant difference and there exists an interaction effect

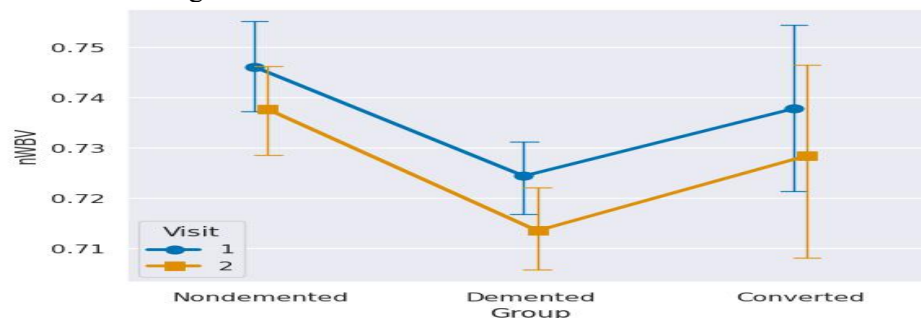


Figure 6: Point Plot of nWBV Across Different Groups and Visits

Table 3: Mixed Effect Anova Table

	Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
0	Group	0.0336	2	141	0.0168	6.7123	1.64e-03	0.0869	NaN
1	Visit	0.0065	1	141	0.0065	94.2512	2.23e-17	0.4006	1.0
2	Interaction	0.0002	2	141	0.0001	1.5335	2.19e-01	0.0213	NaN

Based on the statistical analysis provided in Table 3, the analysis examined the variation in Normalized Whole Brain Volume (nWBV) across different groups and visit periods, as well as the interaction effect between group and visit. Significant differences were found in nWBV among groups ($F = 6.7123$, $p = 1.64e-03 < 0.05$) and across visits ($F = 94.2512$, $p = 2.23e-17 < 0.05$). However, there was no significant interaction effect ($F = 1.5335$, $p = 2.19e-01 > 0.05$).

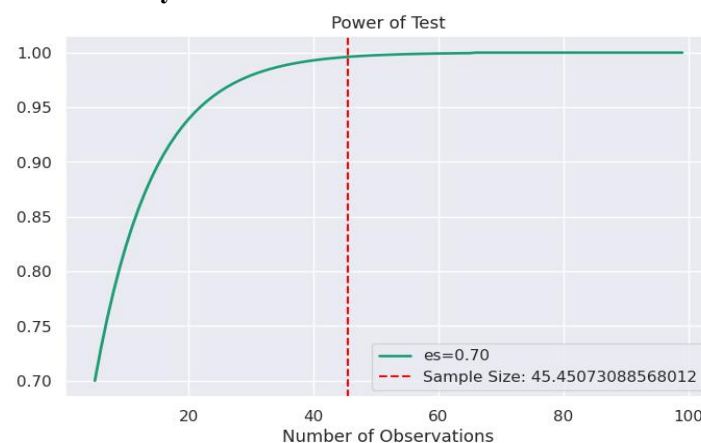
To validate the Mixed Effect ANOVA, certain assumptions need to be met.

Assumption 1: Mauchly's test: P value = 1 > 0.05, assumption of sphericity is met.

Assumption 2: Shapiro-Wilk test: All P values > 0.05, assumption of normality is met.

Assumption 3: Levene's test: P values > 0.05 for both within groups. assumption of homogeneity is met.

Power Analysis:



Based on the findings in Figure 7, it's evident that for a theoretical experiment with a power of 0.91, alpha of 0.05, and an effect size of 0.7, the sample size is approximately 46.

Figure 7: Power Analysis Plot

Conclusion

In summary, this report investigated how dementia impacts eTIV and nWBV across patient groups and visits. Mixed-effects ANOVA analysis showed a significant effect of visits on eTIV, without any interaction. For nWBV, both group and visit time were influential, but their interaction wasn't significant. Assumptions for the ANOVA were met, confirming the validity of the findings. Furthermore, power analysis highlighted the importance of sample and effect sizes for statistical power. These findings deepen our understanding of dementia progression and its effects on brain volumes.